Data Analysis and Decision Making – II
Prof. Raghu Nandan Sengupta
Department of Industrial & Management Engineering
Indian Institute of Technology, Kanpur

Lecture – 42
MAUT

Welcome back my dear friends, a very good morning, good afternoon, good evening to all of you wherever you are, whether in India or outside India. And as you know this is a course which is titled DADM-2 which is Data Analysis and Decision Making II under the NPTEL MOOC series.

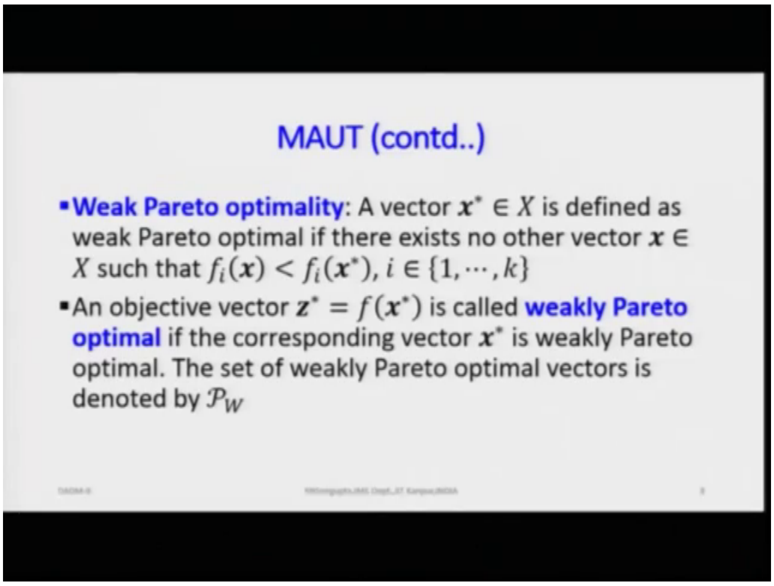
And this total course duration as you all know which I do repeat in the starting of a each class to make you aware, where we are how we are proceeding. This is a 12 week course, total number of lectures is 60 in number, that is 30 hours because each lecture is for half an hour. And as already mentioned and as the plan is we complete 5 lectures each week each again I am repeating is for each of the lecture is for half an hour, and after each week we have assignments.

So, and my good name is Raghunandan sengupta from IME department at IIT Kanpur. So, as you can see in the slide which is the DADM-2 which is number 42 which is in the 9th week; that means, we are in the second class or the second lecture in the 9th week. So, we were discussing about general concepts of domain will we will discuss something to do with dominance, then first order, second order also consider different type of Pareto optimality. I did mentioned about what Pareto optimality in a very simple sense.

Then we will consider the simple properties of MAUT which is Multi Attribute Utility Theory based on utilities are which are based for different attributes and we have multi-criteria's under on a attributes. So, define they are different attributes based on which we find out the utility and try to find out the conglomeration of all of the utility functions based on attributes. Such that we can take a decision whether alternative one is better than two or two is better than three and all so on and so forth.

And we will next; we will all also consider the very simple concept of concept wise what is multi-objective optimisation or multi-criteria decision making. And we will only give you the feel of the problem because these concepts in more details would be solved under the DADM-3 where will consider the main focus would be in the area of optimisation and operation research. So, whenever you are considering MAUT which is Multi-Attribute Utility Theory, we consider the weak Pareto optimality.

(Refer Slide Time: 02:54)



So, consider there are sets a set s x so, x has different elements and each elements or dimension n or whatever the dimension is; that means, if you have x 1 which has an element of x, x 2 which has an element x. Whatever the x value will be we consider the x 1 or x 2 or whichever x is an element of capital X is a vector. And each of these elements in the vectors or values in the vectors correspond with different type of attributes which accrue from any one of the criteria's corresponding to the fact that what is the overall value; net worth you will get by taking that decision x 1 or x 2.

So, the elements of x 1 or elements of x 2 which I have already seen, I am again going to repeat that some of them maybe attributes characteristics which are not easily discernible as quantity values, they have some subjectivity into consideration other part would be the objective criteria which can be quantified. So, let be continue reading it, so if amongst those x 1, x 2 till whichever x value you take if we have some x star.

So, it will be an, and it will be called that x star would be called a weak Pareto optimal if there exist no other vector x from that capital X value, such that the overall functional relationship of say for example, you are taking different type of the functional forms. So, there is in each of the dimensions the functional forms a such the values of x 1 or add values of x 1, x 2, x 3, x 4 is less than equal to that value of the functional form at x star so on and so forth for each and every functional form.

So, if there does not exist any other vector x if this is not true, then we will basically says is weak Pareto optimality would be there; that means, the relation between the x star and the other x's would not be absolute in the sense, there is no dominance on all the aspect, but there would be some dominance or say non-dominance. Such that, the combination how you going to find out which is more dominant and which is less dominant based on which you will rank them may not be clear to you.

So and the objective function which we find out based on the fact when you take the decision of x star or you take the decision x 1 or you take the decision x 2, that will be denoted by z. So, corresponding to x 1 or x 2 or x 3 or x star the corresponding z values would be nomenclature accordingly; that is z 1, z 2, z 3 z star and so on and so forth. So, this would be called the weak Pareto optimum solution. If the corresponding vector x star which I have already find is weakly Pareto optimal and if the set of weakly Pareto optimally solutions can be found out there will be club does a weakly Pareto optimal subset of the capital X.

So, you will basically; technically you should have some sets of x where the overall combinations for different values of the vector values which you have. If the overall utility which I am getting from the utilisation of those particular x's comes out to be the same they would Pareto optimal and in some of the cases, if some of the subsets gives you value for the utility based on those some such sets of x is non-dominating and is less than equal to or is more than equal to, whatever the cases as we change the combination. We will basically have different weak and strong optimality condition considerations based on the Pareto optimal solution.

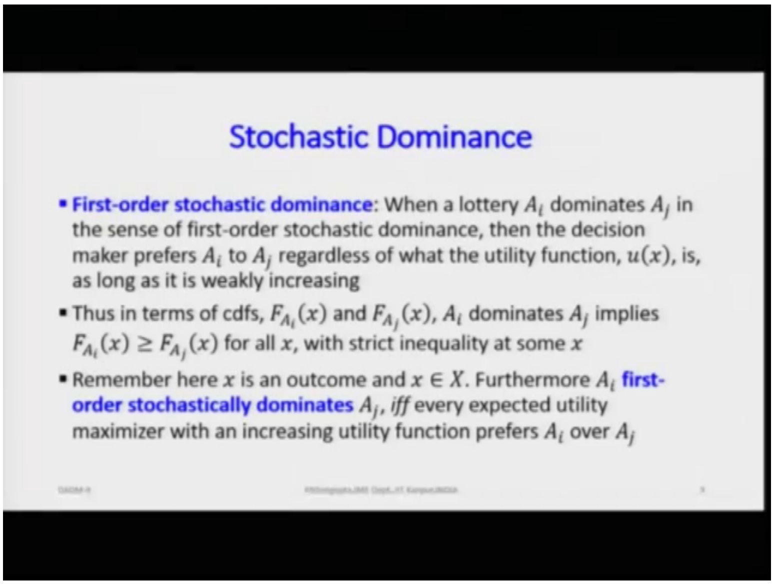
That means, if I have a burn if I have a car and consider the car is brand a, another car is brand b. I am only considered car a and b there can be other brands also. So, consider that brand d is the best. So, in case brand d is the best so, in that case if we consider each and every criteria attributes whatever characteristics you are taking, they would not be the overall brand values, the values which we getting for the brand d would not superseded or would not be more than the any of the individual combination of a, b, c, e, f, g, h whatever you have.

And if they are equal like say for example, d is equal to g in all the accounts for different combinations we will say that they are at they are Pareto optimal solutions are whether you take d or g the overall net value which you which accrues to you as a decision maker are the same. Like you have a budget of 100 rupees as I said and you want to buy different things it can be rice, flour, wheat some vegetables. And if you get then and you are trying to compare them with which respect to say for example, the nutrition values, protein values f or calcium values, carbohydrate values whatever and you have combine them in different proportions.

If you get two basket of such goods based on the concept of funding constraints w which you have 100 rupees as such with overall functional values based on nutrition facts are the same, they would be optimum solutions Pareto optimum solutions such that. Whether you take, one set or one combination of one set of food items with respect to the other set of food items they will give you the same worth.

So, this worth depends on whatever utility function you are trying to use not need not be money, it can be nutrition values, it can be net worth whatever you want to assign. Now, we will consider the concept of stochastic dominance. So, which you have discussed the last day, but I am still repeating it please bear with me

(Refer Slide Time: 08:55)



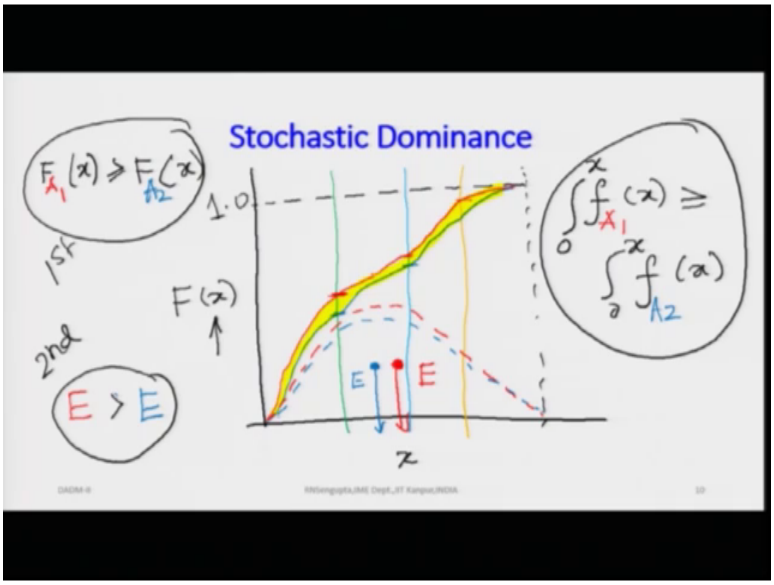
So, the first order stochastic dominance would mean when a lottery. So, lotteries are basically some uncertain decisions, I am considering the lotteries I as A i 1, A i 2, A i 3 for the overall set A.

So, here I mentioning A i and A j, as I did for the TOPSIS electro method, by curve method and all this things. So, when a lottery A i dominates; A j dominance mean the net worth which am getting by utilising the alternative A i or the lottery A i is more than A j. So, it dominates A j in the sense of first order stochastic dominance, then the corresponding decision would be to is prefer A i with respect to A j regardless of what utility function is you are using as long it is weakly increasing. That means, the utility function is weakly increasing and whatever the utility function we use, if it is dash property, then the decision would always be to prefer A i with respect to A j depending on whatever the utility function is then; obviously, we will say it is first order stochastic dominance; that means, A i dominance.

In terms of basically to the CDF values because any of these utility functions would have an f of x; small f of x which is either the PDF or the PMF and correspondingly you will have a capital F of x which is the CDF value. So, the PDF and the PMF are the probability density function which is PDF and PMF is the probability mass function.

So, their corresponding cumulative distribution function would be mentioned as distribution functions or cumulative distribution functions only. And it is denoted the second part is denoted by capital F of x, x is the random variable or the set which you are considering. Thus in terms of CDF, if the functional form of x of F of x based on the fact that you are taking the alternative of lottery A i or a 1 whatever it is because we are mentioning A i so, we will stick to A i. And if the CDF values for the other decision for sad that set of x based on the fact that you are taking the lottery A j. So, we will say that A i dominates A j, if the CDF values for each and every combinations of x F of A i would always be greater than equal to F of A j.

(Refer Slide Time: 11:33)



So, if I am trying to draw a graph; so, let me draw the graph here. So, I am writing let me use the colour black first so F and the case is A, let me mention is A 1. So, it will be easy for me to explain is greater than equal to F of x for A 2. So, this would mean the concept of stochastic dominance which I have mentioned, so let us draw the graph. So, this is the value of 1 because the CDF value always be is less than equal to 1 so, it will go till infinity depending upon the x values.

So, consider I am con see for example, considering some fixed x based on which I will try to draw. So, if consider for; and try to consider this is one value of x, this is another value of x and this is the third value of x, I am taking and here I am drawing F of x and this I am drawing the x values.

So, this value of A 1 would be somewhere here which will dominate, because the sum I am taking the total area for each of them and then if I consider the blue one it is here, similarly if it is this the corresponding values are here and here; that means, the path of the distribution I am drawing trying to draw very simply.

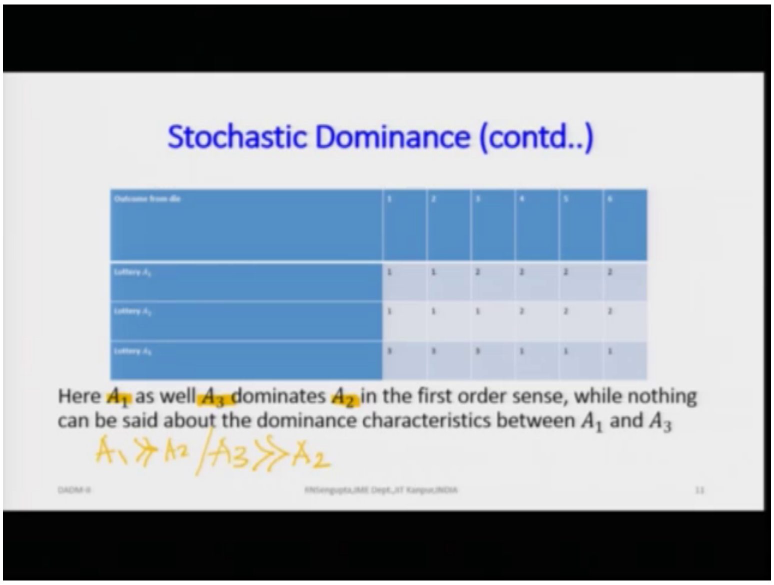
So, considering (Refer Time: 13:45) goes to 1; so, this should be the value of 1. So, this would never exceed, so there is an area such that this would be do you have a different colour let me use the violet; if let me use the yellow would be much lighter so it easy. So, if you see here, the area of the CDF values for any value of x for A 1 and A 2; A 1 would always dominate. So, the area would passed, it is not crossing or it is maybe equal I have not drawn, but it is definitely less. So, in this case why I did not draw the equal 1 is to make you understand.

See if you have this so the CDF values forum case of A 1 would be less, greater than A 2. So, if I write it equational form, so, integration let me take any arbitrary x and it is consider I am starting from minus infinity even though I have drawn it from 0 here. So, 0 some x value f of x for A 1 will be greater than equal to 0 to that corresponding value of x is same, f of x here I mention A 2. So, this being true you will basically have the concept of dominance which I discussed.

So, let us see, thus in terms of CDF values F of x based on the fact that I am taking the lottery A i which was A 1 there and the other lottery which is A 2 there in what I drawn which is A j here. So, if A i or A 1 dominance A j which is A 2; A 2 would imply that the cumulative function distribution for A i or A 1 is greater than A j or A 2. Hence, it is this strictly inequality; with strict inequality at some of the x's, there are some x's where it will dump, it will be greater than some it may be less than equal to. So, remember here x is an outcome so, but for different capital X values there are subsets. So, there are different combinations of x which we have.

So, for some other values is dominates is greater and some of the values its equal to. So, remember here x an outcome and x is an element of capital X for the move A i. First order stochastically dominance A j, if and only if every expected utility maximizer with an increasing utility function would always prefer, A 1 or A i with respect to A 2 which is A j A j, because the corresponding utility which I am getting for that decision would always be more in the former case. So, if I see the graph this is what it portraits.

(Refer Slide Time: 17:09)



So now let us basically take an example. So, along the first row we have the outcome from the die, which is basically 1 to 6 and there are basically 3 lotteries which are correspondingly mentioned in row 2, row 3, row 4 which is lottery A 1, A 2 and A 3. And the corresponding outcomes which you are getting from by rolling the dice, I will only read the row wise for a lottery A 1 the values from 1 to 6 corresponding to the outcomes which I have from the dice 1 to 6 is given as 1 1, 2 2, 2 4. I am mentioning the values listen to it carefully.

So, lottery A 1 again I am repeating would have values 1 for an outcome of 1 in the die, 1 is an outcome of 2 in the die, 2 as an outcome in from the 3 of the die, 2 as an outcome for 4 in the die, 2 as an outcome for 5 in the die, 2 as an outcome for 6 in the die.

Similarly, the corresponding values for lottery A 2 are 1 for 1. The next number which I mentioning is basically the outcome the die, then it is 1 for 2, then it is 1 for 3, then it is 2 for 4; the first number again I am saying is which I mentioning is the outcome. So, here it is 2 out of 4 2 for 4, then it is 2 for 5 and 2 for 6. Similarly, when I have lottery A 3 the corresponding values are 3 for 1, 3 for 2, 3 for 3, 1 for 4, 1 for 5 and 1 for 6.

Now, if I want to consider let us consider one by one. One is strictly the first order dominating strict sense and in the non-strict sense, it is not strictly dominating first order sorry, first order dominance. So, let us consider one and two basically. So, equality holds one and one, one and one; now, if I consider A 1 has started dominating first order because the value is 2. So, if I find out the corresponding CDF values; so, 2 is corresponding to one is more than it is already started dominating. So, 2 is would be more than 1, if I add up all the values from 0 1 0 means 1 till 2.

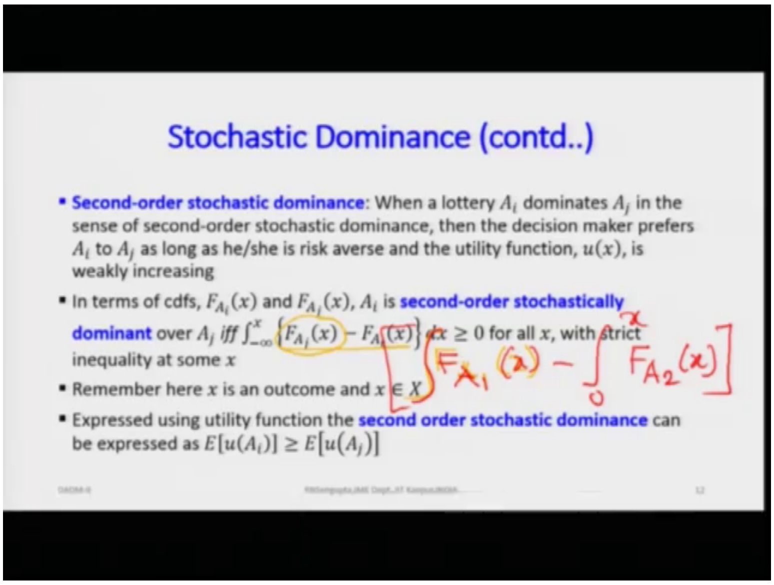
So, the values from 1 to 4 correspondingly it is 1 1 2 2 here it is 1 1. So, equality holds here, equality holds here, it dominance; A 1 dominates first order wise this as the dominances already started. So, this will keep repeating and get much more dominance would become 2, but if the dominance appears in which and every stage individually also. It is equal, fine equal, fine equal to in general, it is greater than equal to for equal to sign for holds 2 for 5 of the cases and greater than sign holds for one of the case.

Now, let us consider A 3 and A 2; so, A 3 it basically means that 3 is better than 1; 3 is better the 1, 3 is better the 1, 1 is here but you have already taken the cumulative value so remember that. So, here is already cumulative value, cumulative value if I take it so; obviously, in the cases the first order stochastic dominance holds. So, if I read it means here, a here A 1 as well as A 3 dominance A 2 in the first order sense. While nothing can be sign about the dominance of A 1 and A 3.

So, let us come why? So, let us first mark so, this dominance this and A 3 also dominance A 2. So, if I want to write, it will be A 1 A 2, another is A 3 A 2. Now, the let us consider the other case of A 3 and A 1. So, if it is A 3 I find out that sum the values we have is weighted (Refer Time: 22:07) the values of A 1 and A A 2. So, it basically is dominating and in the next case it gets reversed 2 is greater than 1. So, and the cumulative distribution form values which I consider would not be coming the same.

So, they would cress cross. So, if I have this one let me go to the slide where I have drawn the diagram. So, if this cream blue one was like this I will erase. So, see for example, it this blue one is what I have done. So, it basically crosses this and then again comes here in this case dominance would not hold to first order because in some of the cases is greater some of the cases it is less.

(Refer Slide Time: 23:06)



Let us consider the second order stochastic dominance. So, in the second order stochastic dominance, we will see a lottery A i, will dominant A j in the sense of second order stochastic dominance then the decision maker prefers A i with respect to A j as long he or she is discovers and the utility function u of x is weakly increasing and that holds true. In terms of the CDF values, a second order stochastic dominant concept would be written.

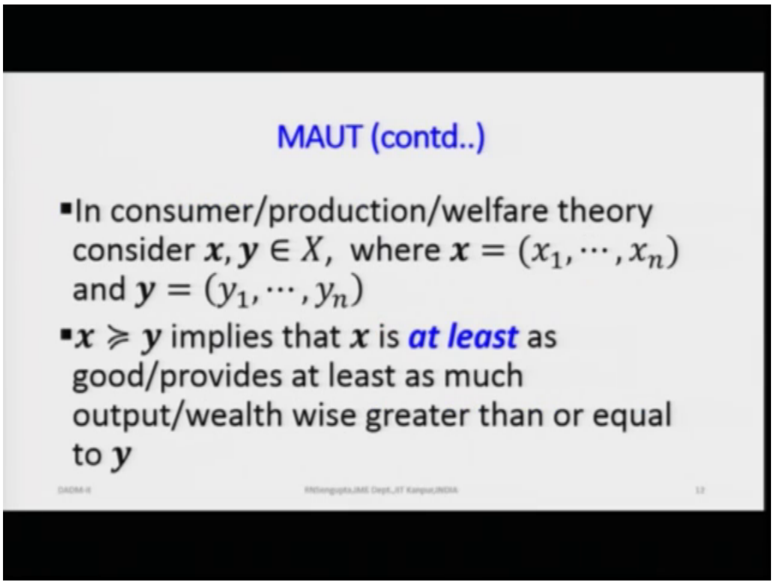
Now, we will take the second integration of the CDF values. So, CDF value was basically the integration of the PDF or the PMF then for the second order stochastic dominance we take the cumulative function of the CDF values already taken.

So, if I consider this the first part I will mark with the orange this part. So, and with the integration tan sin it F of A i would be it should be dominance so; obviously, it should be one i. So, in this case F A i would dominant A j, if the integration so, do not have to do is this would change I will mention it integrating this.

So, if I use the colour say for example, red 0 to same values we can doing for the integration F A 2 x, if this is greater than equal to 0; that means, the cumulative function of A i's each and every step from 0 to x is basically some is greater than the A 2 value. So, I am taking the CDF value, then again integrating and summing up the values. So, in that case this holds true, it is on the second order stochastic dominance.

So, remember here x is an outcome and x is capital small small x is an element of capital X. Express using utility function, the second order stochastic dominance can be expressed as the concept with the expected value for the utility function, A i is greater than the expected value for the utility function considering A j.

(Refer Slide Time: 25:41)



So, to continue the further discussion, in consumer concept consumer concept or the utility which you are getting are from the point of view of production economics or production or welfare excuse me welfare theory.Wwhat is the wealth? The wealth means in the sense of the total welfare with a person not decision maker gets. Considering that if we have 2 distinct sets which are element of capital X, I am denoting this vectors by small x and small y elements are for each x is from or y is from x 1 to x n and for y it is basically from y 1 to y n.

So, if x is better than y, then at least for x at least as much or the net worth which you will get is as good or it will provide the at least the same amount of wealth or wealth wise it would be greater than or equal to y. So if this fact particular holds for the consumer utility basically have at least as a as an amount of good as it accrues to y.

So, x value you will get on the goods front or the net worth front is more than that what you get for the goods are value on for the fact that you utilise the vector y. Similarly, if I come considering from the production concept process production economic. So, it will provide x; will provide at least as much as an output as y it can be more than y also, but at least it will maintain that and from the consider welfare theory. So, the total wealth wise it generated would be more in x at least at to the level of y, so this will hold true.

So, in the first case we consider so let me go to the graph. So, here was the CDF value which is the first order. In the second case, what I will consider is this I will draw this diagram it will be easy for you to understand. So, consider the expected value of the red area is somewhere here, so the PDF you have, let me re draw it at least draw the PDF.

So, consider the PDF is there, I am drawing in a very informal way in order to make you understand through the graph. So, consider the PDF is this, from based on which you get the CDF which is the red line consider this is this, so the green one or the blue one. It is somewhere and also assume the expected value for the red line; the expected values centre of gravity is I am marking as the red colour and the so the blue colour it is here set a gravity.

So, as the expected and this is E. So, if I have this holds true, so this is the second second order and this holds true this should be the first order. So, generally the rule for the stochastic domain which will be this, I will discuss it in later on more in the multi-objective decision making case.

With this I will end this second class in the 9th week and consider more discussions later on accordingly.

Thank you very much and have a nice day.